I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner of Patents and Trademarks, Washington, D.C. 20231, on 9-28-92 (Date of Deposit)

By. Denise Schaefer

CGNE-69

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of)
)
Thompson et al.)
)
Serial No. 07/494,106)
)
Filed: March 16, 1990)
)
For: PLANT STEAROYL-ACP)
DESATURASE - COMPOSITIONS)
AND USES)

Examiner: P. Rhodes

Art Unit: 1804

THOMPSON DECLARATION UNDER 37 C.F.R. 1.132

Honorable Commissioner of
Patents and Trademarks
Washington, DC 20231

Dear Sir:

I Gregory A. Thompson declare as follows:

1. I am a co-inventor named on the above referenced patent application.

2. I have read and understood the McKeon *et al.* references that were cited against the instant application in the Office Action which was mailed on June 22, 1992.

3. The following work was conducted by myself or by others under my direction and supervision. This work demonstrates that the protein preparation described in the McKeon *et al.* references contained a major protein contaminant that interfered with isolation of a desaturase cDNA.

4. Stearoyl-ACP desaturase was purified to apparent homogeneity as described in Example 1 at pages 25-28 of the application. The preparation was evaluated by SDS-PAGE analysis on a 10% acrylamide gel, and a prominent 43 kD band was observed.

5. Protein sequence information was obtained from digestion of the desaturase protein preparation with trypsin and endoproteinase gluC, using methods known in the art. The protein sequence obtained in this manner was used to design synthetic oligonucleotides for use in probing a cDNA bank.

6. The DNA sequence of a clone, 4-4, isolated by hybridization to a synthetic oligonucleotide to the "desaturase" protein sequence was determined. The sequence was compared to an NIH sequence data base by computer aided analysis using the IFIND Sequence Data Bank Searching Program (Intelligenetics, Inc.; Mountain View, CA). Exhibit A documents this analysis. Notebook page 63 (David Shintani notebook S588 013) reports the discovery of homology to a sunflower albumin clone. Pages marked as 2-6 of Exhibit A document the computer search results, the sunflower albumin sequence, HNNG5ALB2, which had the highest matching score, and the DSAT4-4-8 query sequence used in the search.

7. The desaturase preparation was subsequently analyzed by reverse-phase HPLC as described in Example 2 of the application at pages 28-29. The chromatograph resulting from this analysis is shown in Exhibit B (notebook page 75). Notebook page 76 in Exhibit B documents that the smaller peak

which eluted at approximately 29 minutes is designated preDESAT, and the peak eluting later in the gradient (approximately 44 minutes) represents desaturase and is designated DESAT.

8. Amino acid analysis of the protein in the DESAT and preDESAT peaks was conducted, as shown on notebook pages 77 and 83 in Exhibit 2. As shown on page 83 in Exhibit B, the amino acid composition information was used to determine the amount of protein in the preDESAT (39.7µg) and DESAT (64.6µg) peaks.

9. The preDESAT peak was verified as containing the albumin protein contaminant from which amino acid sequence was obtained by comparison of the amino acid composition data discussed above to amino acid composition data from the protein encoded by the isolated albumin cDNA.

10. The amount of albumin and desaturase in the HPLC peaks on a mole basis may be calculated as the amount of protein in the peak divided by the protein molecular mass. The Kórtt *et al.* reference attached hereto as Exhibit C demonstrates that sunflower albumin proteins range in size from approximately 10-18 kD. A similar range is expected for the related safflower albumins, and indeed if the safflower albumin proteins had been larger, they probably would have been detected by SDS-PAGE analysis of the desaturase preparation. Thus, the conservative estimate of 18kD was used as the molecular mass for albumin in determining the molar amount of this protein. The molecular mass of

desaturase used in these calculations is 41.2kD, as determined from the cDNA. These calculations indicate that at least 50% more albumin protein than desaturase protein was present in the "desaturase" preparation on a mole basis.

11. Therefore, the desaturase preparation, which appeared to be homogeneous, was in reality a mixture in which albumin, not desaturase, was the major component on the basis of mole percent, and the presence of this major protein contaminant interfered with isolation of a desaturase cDNA.

DECLARATION

I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Date: 28 September 1992

By: Gregory A. Thompson

- THE RNA SEQUENCING GEL DIDN'T LOOK ANY BETTER THAN IN THE PREVIOUS EXPERIMENT

LAST NIGHT I DISCOVERED THAT MY DSAF CLONE HAS A HIGH DEGREE OF HOMOLOGY TO THE SUNFLOWER ALBUMIN GENE.

TOUGH LUCK EHH?

SUBCLONED PstI SpeI FRAGMENTS FROM S.I.1.1 INTO THE XbaI PstI SITE OF BP10

S.I.1.1

1 5ul H2O + 2ul DNA + 2.2ul 1M 150mM NaCl + 0.5ul 4U/ul PstI + 0.5ul 5U/ul SpeI

16ul H2O + 2ul BP10 DNA + 2.2ul 10X 150mM NaCl + 0.5ul PstI + 0.5ul XbaI

→ 2 HRS 37°C → HEAT KILL @ 65°C → ETHANOL PPT → RESUSPEND IN 16ul H2O + 4ul 5X BRL LIGASE BUFFER + 1ul 0.1U/ul T4 LIGASE. LIGATE O/N AT 14°C.

QUERY = DSAT4-4-8
QUERY-FILE = DSAT4-4.seq, DNA-PARAMETERS,
WINDOW-DNA = 20, WORD-LENGTH-DNA = 4, DENSITY-DNA = LESS
FAST-DNA = YES, GAP-PENALTY-DNA =4

INDEX DIRECTORY
1 nih:

FILE	SEQUENCE	LENGTH	SCORE
(1) hnn.nih	HNNG5ALB2	L=2299	7
2			
(1) yscs.nih	YSCSNF1	L=2587	2
4			
(1) ysp.nih	YSPRSK	L=653	2
2			
(1) phv.nih	PHVHRGPA	L=1113	2
2			
(1) cog.nih	COGCUTA	L=1749	2
2			
(1) ysp.nih	YSPREPDG2	L=3964	2
2			
(1) yscrs.nih	YSCRSCATB	L=212	2
1			
(1) wht.nih	WHTGLIAG3	L=1753	2
0			
(1) yscrn.nih	YSCRNR2	L=2094	2
0			
(1) wht.nih	WHTGLIABB	L=3043	2
0			
(1) wht.nih	WHTGLNB	L=3179	2
0			
(1) wht.nih	WHTGLIABD	L=3310	2

(1) tom.nih	TOMHRGP	L=279	1
(1) bly.nih	BLYB3HORD	L=954	1
(1) wht.nih	WHTGLIA	L=1152	1
(1) wht.nih	WHTGLIABC	L=1156	1
(1) wht.nih	WHTGLU1DG	L=3095	1
(1) bna.nih	BNASSPB	L=713	1
(1) bna.nih	BNANAP	L=718	1
(1) wht.nih	WHTGLIGP	L=798	1
(1) wht.nih	WHTGLGAP	L=2450	1
(1) wht.nih	WHTGLUMPS	L=2915	1
(1) bna.nih	BNANAPA	L=3289	1
(1) brn.nih	BRNWSP	L=621	1
(1) wht.nih	WHTGLIGR1	L=684	1

; DEFINITION Sunflower HaG5 gene for 2 S albumin storage protein.
; LOCUS HNNG5ALB2 2299 BP ds-DNA PLN pre-entry
; ACCESSION X06410.
; SOURCE Helianthus annuus.
; ORGANISM Helianthus annuus; Eukaryota; Planta; Spermatophyta;
; Magnoliopsida; Asteridae;; Asterales; Asteraceae.
; KEYWORDS albumin; storage protein.
; REFERENCE 1 (bases 1 to 2299)
; AUTHORS Allen,R.D., Cohen,E.A., Vonder Haar,R.A., Adams,C.A.,
; Ma,D.P.,, Nessler,C.L., Thomas,T.L. ;
; TITLE Sequence and expression of a gene encoding an albumin storage
; protein in sunflower
; JOURNAL Mol. Gen. Genet. 210, 211-218 (1987)
; STANDARD simple automatic
; COMMENT EMBL features not translated to GenBank features:

;	key	from	to	description
;	PRM	771	774	CAAT box
;	PRM	832	837	TATA box
;	SITE	858	858	CAP site
;	MSG	888	1462	exon 1
;	IVS	1463	1653	intron I
;	SITE	1989	1994	polyA signal

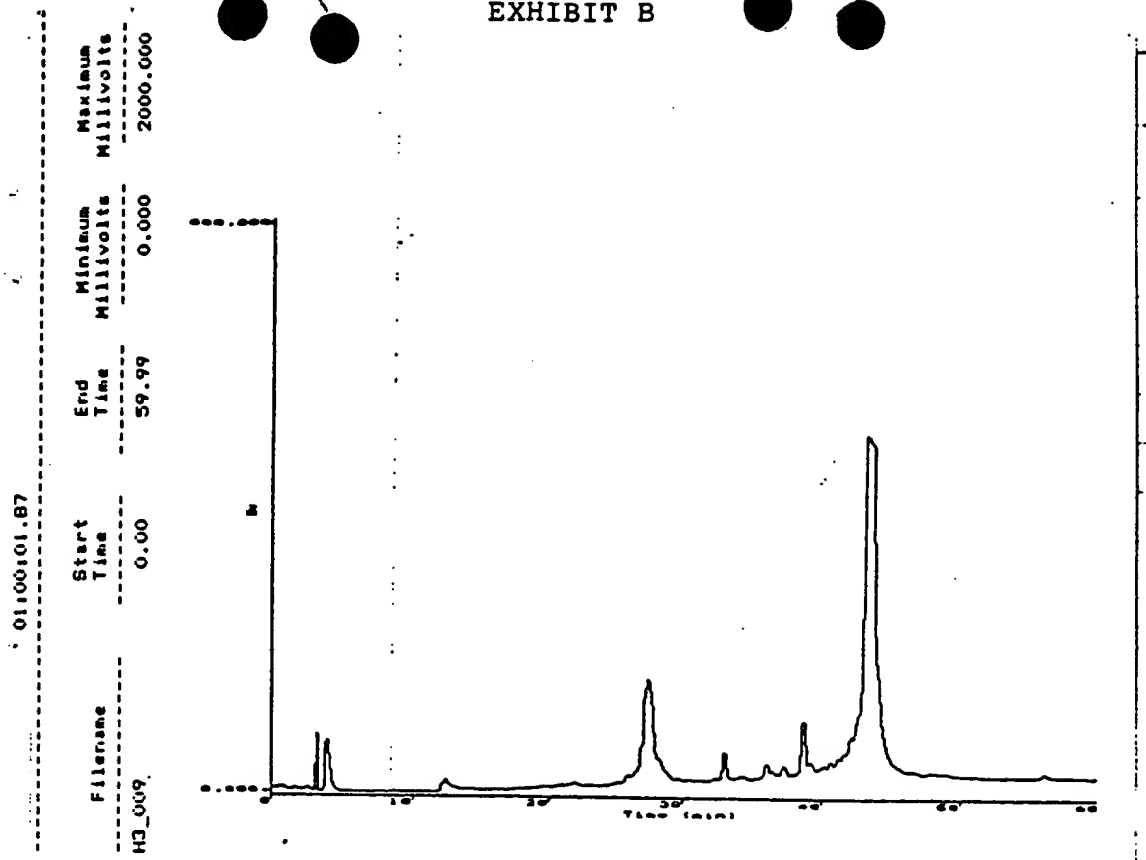
; FEATURES	FROM	TO/SPAN	DESCRIPTION
; pept	888	1462	HaG5 protein
; 1654	1966	HaG5 protein	

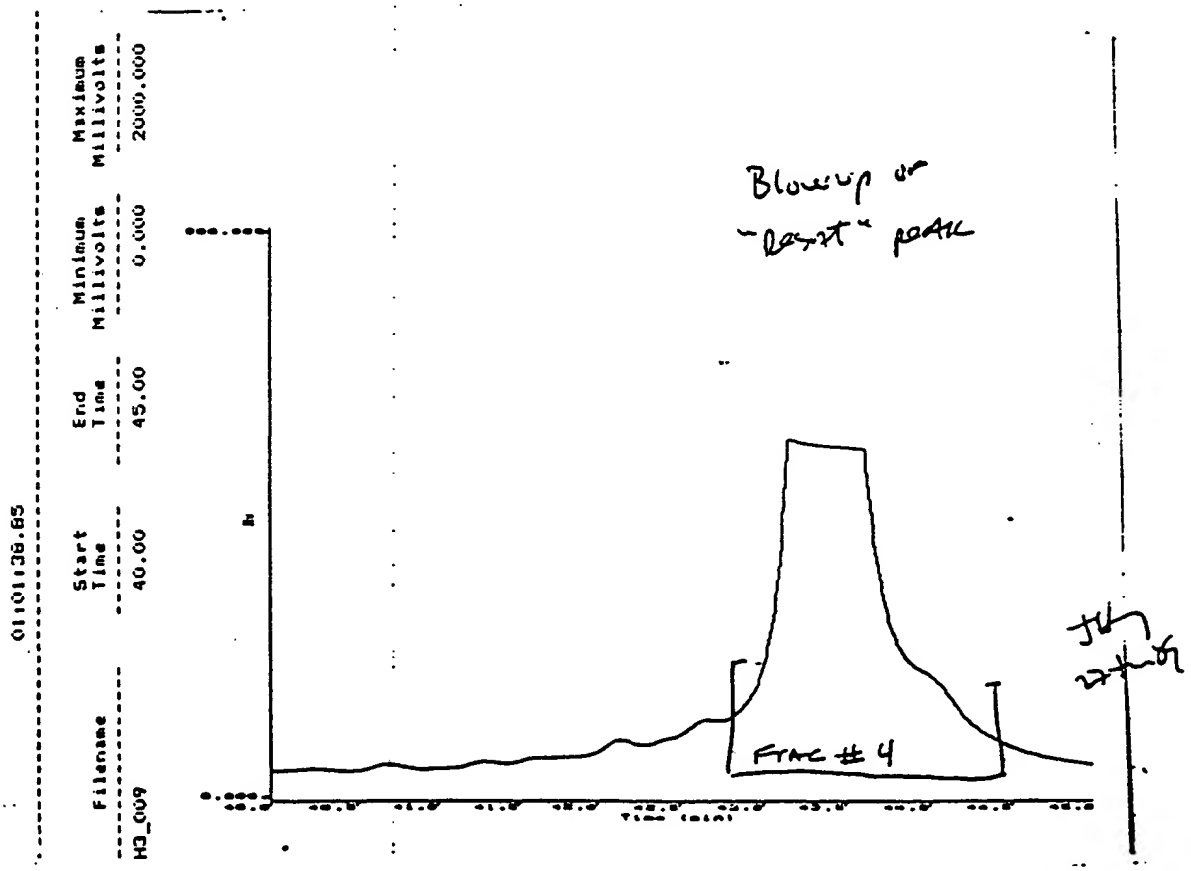
; BASE COUNT 728 a 505 c 453 g 613 t
; ORIGIN
; Found by IFIND in nih:hnn.nih [SHINTANI 22:25]
HNNG5ALB2
gaattctatcactagtgaccaccccatcccctttatttcaataatggaacacaaaaaaaattttaaaaaat
agttgctgttaattgtttaaccgtcattttccaacactttactagctaatcgttaattgatcttcataaa
aaaaaaaattgctatgggtactattgagattgtatatcttatcagttaggcctaaggggcggtcagtga
tattacgaatgatacaaacatcaacgcgtggaacattacaaattcctatccccacctccaagtataacgc
gtgtttgttccacggtttgatgattccgtaatttttttcaacgccgtgatggtttttttttttttttttt
tttgatggtaattgttggttgggggaaattattgggtatggtgttgagtgatgaccacccccactaaaa
aaggttgtgagtgatgtaaaaatggttgctgacatgacgaaacataattggatattgtgagtgataaaat
tttatcattagtgaccaccccgcctcccctttatcatatgttgttatcttccatagttgcggtataccaac
tataggtagtttttatatttatagtttatatttttcattaaactctcttcgccaggctactgtattgtaat
ctataggaatctcaactccacttggaccatccatcatatatttccatttccaaacaaagagaattgacac
ctcatacatactccaaagcatacttccacttgctataattttcatgtaaaaactcgtacgtgttattcga
caatgttcatataacgccaccgattaaactcacctctccacgtatgaacctccacccaccatatatacgc
accaccaccacaccataattcacacaaccacaacaccatctcccacaatggcaaagcaaatagttctcgc
actcgctttcgccgcccttgtagcctttgctaccgcccacacaaccataatcaccaccaccatcgaagac
gagaacccgatctccggacaaaggcaagtgagccaacggatacagggacaaaggctgaaccagtgtcgca
tgttcctccagcagggtcagaacattcctcgcgaattcgataaccctcagatggggcggcagcaggagca
gcagctccagcagtgttgtcaagagctccaaaacatcgaagggcagtgccaatgtgaggcggtgaagcag
gtgttccgagaagcccagcagcaagtacaacagcaacagggacggcagcttgtacccttccgcggttcgc
agcagacccaacagttgaagcagaaggctcagattctccctaacgtatgcaaccttcaatcaagacgatg
tgaaatcggaaccatcaccaccaccgtcaccgagagcaatatcgatatccccttccgtgacaggcccttt
ggcactggatcacaacagtgcagagaaactgaaatccaacgacccgttggtgaatgccaaaggtataata
tactgtaggaaattcataaacatttctttaatatatttatataatgaaaacctacgtattaaatattaat
gaactatatatgttcttttatatataaataaaaaccctactatgtccttagaataatgaaaaccctatgt
attattaatcaacaacatatcttctaatgacgatttccaacaggttcgtggagcagcaaatgcagcagtc
tccgaggtccactagaccataccaacagcggccaggacaacagcagcagcagcagagagggctccaacaa
caatgctgcaacgagctacaaaacgtgaagagggagtgtcattgcgaggcaattcaagaagtggctagga
gagtgatgaggcagccacagcagcagcagcagcaacgtcgtgggcagttcggtgggcaggagatggaaac
cgcgaggagggtgattcagaatctgcccaaccagtgcgacttggaagtccagcaatgcacaacctgtacg
ggatgatgtgaagtatgaaggagctataaataaaggagtggcatctctagtcgttggaggtctttgactt
tgagatgcagggagggtagtttctcctttctaataataataagaataataatgacttgatgagccacaaa

tgtgtacttatagctcatgtgtaatcgttgctttcatgttatttttgatatgttgaatggcacaaatgag
tgatttaatgcccatgaaaaatctcaaacttaggttcttaaagttgattaatgcccaaaaaaaaactaaaa
agtgttgaatttaatgtcaaaattcaatggataattctagccttagaaataagtttaaa1

; DSAT4-4-8 Written [ROSE 8:57]
DSAT4-4-8
CACCACCATCACCACCACCACCACAAACCATCCAAAATGGCAAAGCTCATAGTCCTGGCCTTCACCATCG
CCACCCTCCTAGCCCTCGCCTCCGCGCACCGAACCATCGTCACCACCACCATCGAAGAAGACGACACCAC
CAACCCTAGGTCGCAACGAGAACAACAACAACAGTGCGGGCGACATCTCCAGGGGCAGCAACTCCGGCAA
TGCCAGACTCACCTCCAACAGCCGGATCAAACCCAACAACAAACCCTCCAACAGTGCTGCCAAGAGCTCC
GAAACGTCGAGCAGCAGTGCCAGTGTAGGGCCGTCAAAAAGATATTCGGAGAGGTTGCTGAGCAACAACA
GCAGCAACAAGAAACTGGACCTTTTGGTTCACAGCAGATCCAAAGGCTGAAACAAAGAGCCGANNNCTCC
CAAACCAATNCAACCTCCAAAGCGGAAGACAGTGTGGATTAGGAGCGGGGGAGAGTNCCGGCAGAGGTCC
AAGGANGGCAGTTTAA1

EXHIBIT B





Adjust pH to 7-8 with NH4OH on Both FRAC 4 Composite & FRAC 5 Composite.

Neal P. Dr

Analyses of HPLC'ed DESAT - gels

vol in peak 5 = 1.37 mls (small peak @ mid grad)
vol in peak 4 = 1.10 mls (DESAT peak @ late grad).

Rinse orange-capped tubes c̄ HPLC H_2O & combine, add to stock to make vol =

pre DESAT peak — 1.5 mls
DESAT peak — 1.3 mls.

① run PHAST SDS gel of above vols.
run " " " of DEAE → 48-F52 (stored frozen), plus DEAE/APS pools.

~~3λ stock + 3λ LSB + pre → 15' @ 65° → SDS 8-25 phast gel.~~
↓
~~5 samples~~

#1 stocks APS, DEAE, F48 ↓ F52 } 3λ + 3λ LSB + 1% pme → 65° 5' → 8-25 phast

#2 pre Desat, DESAT } 3λ + 3λ LSB
pre Desat, Desat } 3λ + 3λ LSB + 1% pme
} 65° 5' → ~~8-25~~ 10-15 phast

post HPLC pool of F44-F52 } 3λ + 3λ LSB + 1% pme, 3λ + 3λ LSB → 65° 5' → 10-15 phast

Christopher B. Glascock

JJL

Analysis of HPLC'd resat - AAA

λ/ hydrolysis tube

1. pre Resat peak ; Frac #5 — 50λ
2. Resat peak ; Frac #4 — 50λ
3. pool F48-F52, pre HPLC — 50λ
4. B/A @ 20 ug/ml w/v — 50λ
   50λ @ 1/10 dil in H₂O.
5. H₂O
6. 50% B/A buffer — 50λ
7. Blank tube

↓

Speedvac to dry

↓

110°, 24h in vacuo + phenol.

JH

HCl Amino acid analysis anal of _ pre[illegible]C DESAT
Beckman 6300
Chrom # CA022_008 inj vol. 50
hydrol vol. 65

amino acid	pmoles/ inj vol	ngm/ inj vol	mole %
ASX	320.5	36.73	10.6%
THR	191.6	19.37	6.4%
SER	179.8	15.66	6.0%
GLX	511.3	65.80	17.0%
GLY	366.5	20.93	12.2%
ALA	231.9	16.49	7.7%
VAL	200.2	19.85	6.6%
CYS	0	0.00	0.0%
MET	62.72	8.22	2.1%
ILE	138.2	15.64	4.6%
LEU	241.9	27.38	8.0%
TYR	86.78	14.16	2.9%
PHE	114.5	16.85	3.8%
HIS	67.45	9.25	2.2%
LYS	149.6	24.64	5.0%
TRYP	0	0.00	0.0%
ARG	151.3	23.63	5.0%
PRO	0	0.00	0.0%
	3014.25	334.6	100.1%

μg/inj = 0.3346
μgm hydrol = 0.4350

76.17, ÷ 50 = 38.5   1,000 / 38.5 = 511.5   (113 μg)

NL = 35734

HCl Amino acid analysis anal of _ preDESAT
Beckman 6300
Chrom # CA022_006 inj vol. 50
hydrol vol. 65

amino acid	pmoles/ inj vol	ngm/ inj vol	mole %
ASX	527.8	60.49	7.6%
THR	362.1	36.61	5.2%
SER	621	54.09	8.9%
GLX	1807	232.56	25.9%
GLY	799.8	45.67	11.5%
ALA	324.8	23.09	4.7%
VAL	452.9	44.91	6.5%
CYS	0	0.00	0.0%
MET	91.6	12.01	1.3%
ILE	209.1	23.67	3.0%
LEU	607.6	68.78	8.7%
TYR	158.9	25.93	2.3%
PHE	201.2	29.62	2.9%
HIS	91.61	12.57	1.3%
LYS	188.6	31.06	2.7%
TRYP	0	0.00	0.0%
ARG	524.7	81.96	7.5%
PRO	0	0.00	0.0%
	6968.71	783.02	100.0%

1.0 μg / 38.5 λ

μg/inj = 0.7830
μgm hydrol = 1.0179

$^{5}/_{65}$ = 50 from 1500
38.5

38.96 × 1.0179 = (39.7 μg)

76.17, ÷ 50 = 38.5

NL = 35633   97.1

HCl Amino acid analysis anal of _ DESAT
Beckman 6300
Chrom # CA022_007 inj vol. 50
hydrol vol. 65

amino acid	pmoles/ inj vol	ngm/ inj vol	mole %
ASX	1334	152.88	10.3%
THR	893.9	90.37	6.9%
SER	683.8	59.56	5.3%
GLX	1830	235.52	14.2%
GLY	1036	59.16	8.0%
ALA	1068	75.93	8.3%
VAL	1010	100.14	7.8%
CYS	0	0.00	0.0%
MET	351.9	46.13	2.7%
ILE	642.9	72.78	5.0%
LEU	1182	133.80	9.2%
TYR	410.6	67.01	3.2%
PHE	579.9	85.36	4.5%
HIS	442.1	60.66	3.4%
LYS	726.1	119.59	5.6%
TRYP	0	0.00	0.0%
ARG	721.2	112.65	5.6%
PRO	0	0.00	0.0%
	12912.4	1471.54	100.0%

1.9 μg / 38.5 λ

μg/inj = 1.4715
μgm hydrol = 1.9130 = 64.6 μg

31.8.
76.17 ÷ 50 = 38.5   1200 / 38.5 = 33.8

33.8 × 1.913 = (64.6 μg)

NL = 35807   1003

HCl Amino acid analysis anal of _ 50% buffer A
Beckman 6300
Chrom # CA022_012 inj vol. 50
hydrol vol. 65

amino acid	pmoles/ inj vol	ngm/ inj vol	mole %
ASX	30.43	3.49	7.5%
THR	14.04	1.42	3.5%
SER	43.66	3.80	10.8%
GLX	45.6	5.87	11.3%
GLY	87.54	5.00	21.6%
ALA	27.62	1.96	6.8%
VAL	0	0.00	0.0%
CYS	0	0.00	0.0%
MET	0	0.00	0.0%
ILE	32.75	3.71	8.1%
LEU	27.04	3.06	6.7%
TYR	0	0.00	0.0%
PHE	0	0.00	0.0%
HIS	0	0.00	0.0%
LYS	96.2	15.84	23.8%
TRYP	0	0.00	0.0%
ARG	0	0.00	0.0%
PRO	0	0.00	0.0%
	404.88	44.15	100.1%

μg/inj = 0.0442
μgm hydrol = 0.0575   = 3% DESAT
= 5.6% preDESAT

NL = 35625

CBlood   JH

Phytochemistry, Vol. 29, No. 9, pp. 2805–2810, 1990.
Printed in Great Britain.

0031-9422/90 $3.00 + 0.00

EXHIBIT C

LOW MOLECULAR WEIGHT ALBUMINS FROM SUNFLOWER SEED: IDENTIFICATION OF A METHIONINE-RICH ALBUMIN

ALEXANDER A. KORTT and J. BRUCE CALDWELL

CSIRO, Division of Biotechnology, Parkville Laboratory, Melbourne, Australia

(*Received* 30 *January* 1990)

Key Word Index—*Helianthus annuus*; Compositae; sunflower albumins; amino acid compositions; methionine-rich albumin.

Abstract—The small M_r proteins of sunflower seed (*Helianthus annuus*) are soluble in 60% (by vol) methanol. These proteins, classified as albumins on the basis of their solubility in water, were isolated from a salt extract of sunflower seed by precipitating the 11S globulins with 60% (by vol) methanol and were resolved into eight distinct components by reversed-phase HPLC. Electrophoresis showed that each fraction contained a single polypeptide chain with an apparent M_r in the range 10 000–18 000. The individual sunflower albumins are basic proteins with distinct amino acid compositions. The major albumins (4–8) contain high contents of glutamine/glutamic acid, asparagine/aspartic acid, arginine and cysteine, characteristic of the 2S class of seed storage proteins. One exception was the small glutamine/glutamic acid content of albumin 6. Two of the sunflower albumins (7 and 8) with $M_r \sim 10\,000$ were methionine-rich proteins containing 16 residues per cent methionine as well as eight residues per cent cysteine. These sulphur-rich proteins constitute some 7% of the total salt extractable seed protein. A method for the preparation of these two albumins using a reversed-phase Sep-pak cartridge is described.

INTRODUCTION

Storage proteins are major components of seeds and in general are characterized by a high percentage of glutamine which provides a source of nitrogen for the developing seedling [1]. The seeds of most dicotyledonous plants contain two major protein classes, globulins and albumins, which are distinguished on the basis of solubility [2]. The 7S and 11S globulins of nutritionally important legumes and oilseeds have been extensively studied. The albumins are a more diverse group and are usually classified as '2S' proteins. Recently a family of 2S proteins has been recognized whose major role appears to be that of storage proteins [3]. These small M_r albumins which are rich in cysteine, arginine, glutamine and asparagine, have been reported to occur in many species of oilseeds [3]. They are devoid of protease inhibitory activity, but some are reported to be allergens [4, 5]. The major albumins from several seeds, including castor bean [6], rapeseed [7], yellow mustard [8] and Brazil nut [9] have been characterized and shown to be structurally related [10–13]. This family of albumins is related to the wheat 2S albumins and CM-proteins, and inhibitors of trypsin and α-amylase from cereals [14].

Sunflower seed contains two major protein classes having sedimentation coefficients of *ca* 11S and 2S, with minor species of 7.8S and 18.1S [15–17]. The major 11S globulin (helianthinin), an oligomeric protein of $M_r \sim 300\,000$ with a subunit structure similar to that of other 11S globulins like pea legumin, has been studied extensively [18].

The sunflower 2S albumin fraction, which constitutes ~20% of the total seed protein, is basic in nature and rich in cysteine with a high content of α-helix structure [19]. Fractionation on Sephadex G-75 and CM-Sephadex C-50 showed that the sunflower seed albumin fraction contained at least two components [20, 21], which were not further characterized. Because the sunflower albumins contribute to the nutritional value of sunflower seed meals and protein isolates, it was of interest to characterize this protein fraction and to establish its relationship to other 2S proteins recently described. In this paper we describe a method for the isolation of the sunflower seed albumins and their fractionation by reversed-phase HPLC into eight distinct components. Amino acid analysis of these components revealed the presence of two sulphur-rich albumins containing *ca* eight residues per cent cysteine and 16 residues per cent methionine. A convenient method for the preparation of the major sulphur-rich albumin is described.

RESULTS

Isolation of the sunflower albumins

SDS-PAGE analysis of sunflower seed proteins extracted with buffered SDS or buffered 0.5 M NaCl at pH 8 showed two groups of polypeptides (Fig. 1). The major group of larger M_r polypeptides (M_r 50 000–65 000), which are disulphide-linked, corresponds to the subunits of helianthinin [18] the 11S globulin of sunflower. The less abundant group of smaller M_r polypeptides (M_r 10 000–14 000) represents the 2S albumins. The addition of methanol (60% v/v) to a buffered salt extract quantitatively precipitated the globulins and other minor polypeptides (Fig. 1a, lane 3) but not the albumins which were recovered by subsequent acetone precipitation (Fig. 1a,

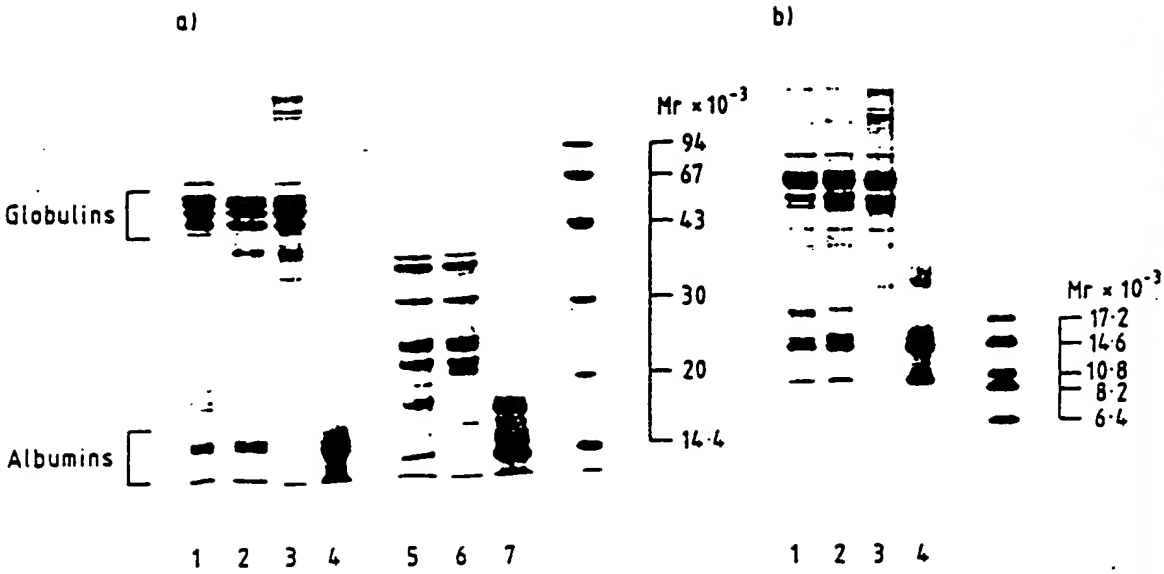


Fig. 1. SDS-PAGE of sunflower seed proteins. (a) Samples analysed using the Laemmli system in the absence of 2-mercaptoethanol: Lane 1, buffered SDS extract of sunflower seed meal; lane 2, buffered salt extract of sunflower seed meal; lane 3, globulin fraction; lane 4, albumin fraction (see Experimental). Lanes 5–7 are the same samples as in lanes 2–4 but run in the presence of 2-mercaptoethanol. (b) Samples as in lanes 1–4 in Fig. 1a analysed using the Tricine gel system [33] in the absence of 2-mercaptoethanol. The M_r values of the standards in the PMW kit are the values reported [33].

lane 4). Electrophoretic analysis showed that the sunflower albumin fraction contained several components. SDS-PAGE, using the Tricine buffer system, showed two to three major components with apparent M_rs of 14000 and ~10000 (Fig. 1b). PAGE under non-denaturing conditions at pH 4.3 also showed that the albumin fraction contained a number of distinct proteins (see Fig. 4).

Chromatography of the albumin fraction on a Waters μBondapak column resolved eight discrete peaks referred to as sunflower albumins (SFA) 1 to 8 according to their order of elution as indicated in Fig. 2. Peaks 1 to 8, pooled from several preparative runs, were further purified by rechromatography on a Vydac 218TP54 column prior to electrophoretic and amino acid analyses. SFA's 1, 3–6 and 8 each yielded essentially a single major peak on rechromatography, while SFA 2 and SFA 7 yielded two peaks (a and b) on the Vydac column. SFA 7 and SFA 8, the sulphur-rich albumins (see below), were separated in mg quantities from the other albumins on a Sep-pak C_{18} cartridge using conditions which selectively bound SFA 7 and SFA 8. The results of the Sep-pak separation are shown in Fig. 3; with the Sep-pak equilibrated with 50% solvent A/50% solvent B SFA's 1–6 did not bind (Fig. 3a), and bound SFA 7 and SFA 8 were eluted with 40% solvent A/60% solvent B (Fig. 3b). The fraction containing SFA 7 and SFA 8 was then chromatographed on the Waters μBondapak column to yield SFA 7 (Fig. 3c) and SFA 8 (Fig. 3d).

Electrophoretic analysis

Electrophoretic analysis at pH 8.8 showed that the sunflower albumins did not enter the gel and that they were basic proteins (pI > 8.8) as reported previously [19]. At pH 4.3 each albumin component separated by RP-HPLC migrated essentially as a single discrete protein band (Fig. 4). Notably, SFA 3 (data not shown) and SFA 6 (Fig. 4) were more basic than the other albumins.

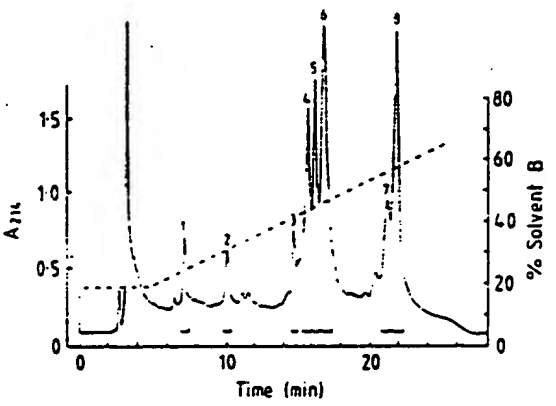


Fig. 2. Reversed-phase chromatography of sunflower seed albumins on a μBondapak column (Waters) with a linear gradient of acetonitrile in 0.1% (v/v) TFA from 20 to 65% solvent B in 20 min. The flow rate was 1 ml min^{-1}. The peaks were collected as indicated by the bars.

SDS-PAGE analysis, in the absence and presence of 2-mercaptoethanol, yielded a single subunit with a characteristic apparent M_r for each albumin, except SFA 3 which contains a second minor band (Fig. 5). Reduction showed that the sunflower albumins are composed of a single polypeptide chain and not two disulphide linked polypeptide chains as found for the 2S albumins of other seeds such as Brazil nut [13]. The isolated sunflower albumins fall into three distinct size classes (SFA's 1–2, M_r ~18000; SFA's 3–6, M_r ~14000; SFA's 7–8, M_r ~10000) consistent with the assignment of the albumin bands in Fig. 1.

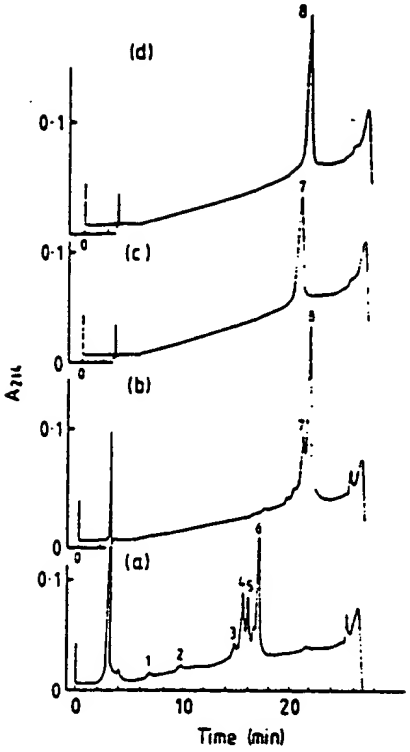


Fig. 3. Analytical reversed-phase chromatography on a μBondapak column (Waters) as described in Fig. 2 of the albumin fractions separated by the Sep-pak cartridge (see Experimental) and of purified SFA 7 and SFA 8. (a) albumin fraction not bound to the Sep-pak. (b) albumin fraction eluted from the Sep-pak with 60% solvent B. (c) SFA 7 and (d) SFA 8 after preparative separation on the μBondapak column.

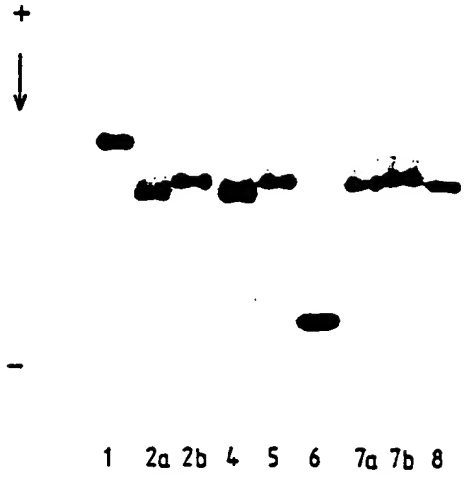


Fig. 4. Electrophoretic analysis of the purified sunflower albumins at pH 4.3. The lane numbers correspond to the respective albumins separated by reversed-phase HPLC as shown in Fig. 2 and in the text.

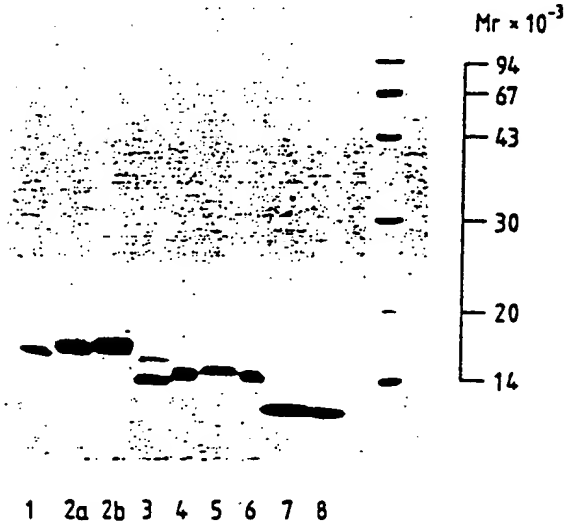


Fig. 5. SDS-PAGE of the purified sunflower albumins in the absence of 2-mercaptoethanol using the Laemmli system. The lane numbers correspond to the respective albumins separated by reversed-phase HPLC as shown in Fig. 2 and in the text.

Amino acid compositions

The amino acid composition of the total sunflower albumin fraction was similar to that reported previously [19, 20]. The compositions of the eight albumins separated by RP-HPLC are presented in Table 1.

As a group, the sunflower albumins are characterized by high contents of Glx, Asx and Arg, and a ½Cys content of four to eight residues per cent, a feature shared with the 2S proteins from various seeds [3]. The individual sunflower albumins, however, show some interesting differences in their amino acid compositions. For example, the minor components SFA 1 and SFA 3 contain no ½Cys residues and SFA 1 contains no His, Phe or Trp but is rich in Glx (27%) and Gly (15.8%) residues. The two components of SFA 2, a and b, although similar in composition are characterized by differences in His and Phe levels (Table 1). SFA 3 is characterized by relatively smaller contents of Lys (4.9%) and Glx (12.7%) and relatively larger contents of His (4.4%), Ser (10.8%), Val (4.4%) and Ile (4.1%) than SFA's 1 and 2. The ratio of the A_{275} at 275 nm to 290 nm indicated the presence of only Tyr residues in SFA's 1, 2 and 3.

SFA 4 and SFA 5 are characterized by a large content of Glx and Arg, and spectral analysis indicated that SFA 4 contained no Tyr or Trp residues while SFA 5 contained only Tyr residues. In contrast, SFA 6, one of the most basic albumins, contains only 2.5% Glx, lacks His, Phe and Trp residues and is relatively rich in Asx, Thr, Pro and Ala residues. The amino acid compositions of the two smaller M_r sunflower albumins, SFA 7 and SFA 8, are characterized by a large content of sulphur containing amino acids with 16 residues per cent Met and 8 residues per cent ½Cys (Table 1). The compositions of SFA 7 and SFA 8 are identical and the two components of SFA 7 (a and b) separated on the Vydac 218TP54 column had the same composition (data not shown) as the parent material. Spectral analysis showed that SFA 7 and SFA 8 contained both Tyr and Trp residues and sequence data [22] showed a single Trp residue in SFA 8. Both SFA 7 and SFA 8 showed the same mobility on SDS-PAGE (Fig. 5) but a small difference in mobility was

Table 1. Amino acid composition of sunflower seed albumins

	Residues per 100 residues* Albumins									Total albumin fraction	
	1	2a	2b	3	4	5	6	7	8	†	‡
Lys	2.1	10.4	12.4	4.9	2.9	1.0	8.0	4.1	4.0	5.2	5.3
His	0.2	0	1.1	4.4	0.9	0.8	0	2.9	2.9	1.9	1.5
Arg	7.6	6.5	6.5	7.0	8.1	11.4	5.6	6.6	6.7	7.8	5.2
Asx	5.9	7.1	7.1	8.2	4.7	6.6	10.1	8.8	8.6	7.8	7.4
Thr	7.5	2.8	2.3	6.0	3.0	1.6	11.0	1.0	1.0	4.1	4.8
Ser	4.0	8.8	7.2	10.8	4.9	3.1	5.3	4.6	4.5	4.9	5.9
Glx	27.2	17.8	19.5	12.7	38.1	31.3	2.5	19.4	19.3	20.6	18.4
Pro	2.4	3.9	4.0	7.1	2.9	7.9	11.6	5.8	5.6	6.3	3.9
Gly	15.8	17.8	17.7	13.0	4.8	5.0	7.7	5.1	4.9	8.3	14.2
Ala	7.1	4.2	3.3	6.2	3.0	2.7	10.9	3.1	3.0	5.4	5.1
½Cys	0	4.0	4.4	0	5.6	5.6	8.1	6.9	6.7	6.6	6.6
Val	2.2	3.5	2.9	4.4	5.4	6.1	3.5	2.1	1.9	4.1	5.1
Met	1.1	1.3	1.7	1.2	0.9	3.2	1.2	14.7	15.0	4.5	2.8
Ile	1.2	1.9	1.1	4.1	3.5	5.2	5.5	3.0	2.9	3.6	4.0
Leu	4.7	4.0	3.6	6.4	8.5	3.7	7.9	9.1	9.0	6.0	5.8
Tyr	1.3	4.0	4.6	1.9	0	2.3	1.2	2.9	3.0	1.8	1.9
Phe	0	1.9	0.9	1.9	2.9	2.5	0	0	0	1.5	2.0
Trp	0	n.d.	n.d.	n.d.	0	n.d.	n.d.	n.d.	1.0	n.d.	—

*The values are from 24 hr hydrolysates and are uncorrected.
†Albumin fraction isolated after precipitation of globulins with 60% (v/v) methanol (see Experimental).
‡Data of Youle and Huang [3].

apparent on PAGE at pH 4.3 (Fig. 4) suggesting a charge difference due perhaps to amide differences. The amino-terminal sequences of SFA 7 and SFA 8 (PYGRGRT) were identical.

DISCUSSION

The small M_r protein fraction of sunflower seed, which is soluble in 60% (v/v) methanol, contains eight distinct components which can be readily separated and purified by reversed-phase HPLC. Previous purification methods [20, 21] indicated the presence of only two components. RP-HPLC therefore provides a greatly improved method for resolving individual seed proteins compared with previous fractionation methods. For example, the 2S protein fraction of linseed was resolved into six distinct components by RP-HPLC (Kortt and Caldwell, unpublished data).

The albumin fraction of sunflower seed contains four major (SFA's 4–6 and 8) and four minor (SFA's 1–3 and 7) proteins. Six of the sunflower albumins (SFA's 2, 4–8) are characterized by a large content of cysteine (5–8 residues per cent), glutamic acid, aspartic acid and arginine, and in this regard they are similar to the 2S proteins found in various seeds [3]. The individual sunflower albumins show some marked differences in amino acid composition, the most notable being the absence of cysteine in SFA 1 and SFA 3 and the unusually high methionine content in SFA 7 and SFA 8. These methionine-rich proteins in sunflower represent *ca* 37% of the total albumin fraction and *ca* 7% of the total sunflower protein.

Comparison of the amino acid compositions of 2S proteins from various seeds [3] including cotton, linseed, lupin, hazel nut, Brazil nut, rapeseed, castor bean and sunflower indicated only one species, Brazil nut, with unusually high levels of sulphur amino acids, in particular methionine. Recently, the 2S albumin from Brazil nut, which accounts for some 30% of the total seed protein, was isolated and shown to be a sulphur-rich protein with 18 residues per cent methionine and eight residues per cent cysteine. The isolation of a methionine-rich 2S protein from sunflower seed suggests that methionine-rich 2S proteins may occur also in other dicotyledonous seeds albeit in smaller amounts. Methionine-rich proteins have been isolated from maize [23, 24] with a M_r 10000 zein containing 21 residues per cent methionine and three residues per cent cysteine [25]. These zeins, however, are not related structurally [24, 25] to the methionine-rich Brazil nut 2S protein [15]. Structural studies [7, 10, 11, 13] have shown that the two-chain 2S proteins from the various seeds are related, and they are characterized by the conservation of the number and position of the cysteine residues. These dicotyledonous 2S seed proteins are also related to the 2S albumins of wheat and inhibitors of trypsin and α-amylase from cereals [14] which are single-chain polypeptides. In contrast to the other dicotyledonous 2S proteins, the sunflower albumins are single-chain polypeptides. Sequence studies [22] have shown that SFA 8 is related to the other seed 2S proteins, and in particular to the methionine-rich Brazil nut 2S protein. Thus SFA 8 also belongs to the family of 2S seed proteins described by Kreis *et al.* [14].

As many of the major plant proteins, especially those of legume seeds, are deficient in the essential sulphur amino acids, the discovery of a single polypeptide chain rich in cysteine and methionine raises the possibility that such a protein may be a useful vehicle for improving the nutritive value of plant proteins by genetic engineering. The introduction into vegetative tissue of pasture plants of a sulphur-rich protein which also is resistant to degradation in the rumen is of particular interest [26], as it has been demonstrated that an increase in the supply of sulphur-amino acids such as methionine in sheep's diet significantly increases the growth of wool [27]. Spencer

et al. [28] have identified pea and sunflower proteins which are resistant to rumen fluid degradation *in vitro*. The sunflower albumins were resistant to rumen degradation, and the methionine-rich albumins SFA 7 and SFA 8 were particularly stable under these conditions. These results suggest that SFA 8 is most suitable for incorporation into transgenic plants.

EXPERIMENTAL

Materials. Seeds of *Helianthus annuus* L. cv Hysun were obtained from the Pacific Seed Co. The sources of other materials were as previously described [29].

Isolation of the albumin fraction. Dry sunflower seeds were ground in a Retsch cyclone mill (0.75 mm mesh), defatted with petrol and air-dried. Defatted meal (10 g) was extracted with 0.02 M TES buffer, pH 7.8, containing 0.5 M NaCl–1 mM PMSF (500 ml) for 2 hr at room temp. The slurry was filtered through terylene cloth and centrifuged at 13 500 *g* for 30 min. The clarified extract was cooled to 0° and MeOH was added to 60% (v/v) (750 ml) to ppt. the globulins quantitatively. The ppt. was recovered by centrifugation (13 500 *g*, 30 min) and the albumins were precipitated from the supernatant with Me_2CO (5.5 l) at −20° and recovered by centrifugation. The pptd albumins were dissolved in H_2O and dialysed extensively against H_2O. The pigmented residue obtained on dialysis was removed by centrifugation, and the soluble proteins were recovered by lyophilization. The yield was 20 mg g^{-1} of defatted seed meal.

Fractionation of the albumins. The albumin fraction (5 mg ml^{-1}) was dissolved in 0.1% (v/v) TFA and centrifuged to remove insoluble material. RP-HPLC was carried out at 45° on a Waters μBondapak column (4.6 × 300 mm) or a Vydac 218TP54 column (4.6 × 250 mm) using a TFA–acetonitrile gradient. Solvent A was 0.1% (v/v) TFA and solvent B was 0.1% (v/v) TFA–70% (v/v) acetonitrile. All solns were filtered through a 0.2 μm filter (Magna Nylon 66) and degassed prior to chromatography. The albumin components separated on the Waters μBondapak column were collected manually and the fractions from several runs were pooled and dried under vacuum at 50°. The individual components were rechromatographed on the Vydac 218TP54 column prior to electrophoresis and amino acid analysis.

Preparation of the sulphur-rich albumins. The sulphur-rich albumins (SFA 7 and SFA 8) were separated from the other albumins in a single step using a Sep-pak C_{18} cartridge (Waters-Millipore). The Sep-pak cartridge was equilibrated with 50% solvent A–50% solvent B and the albumin fraction (25 mg) was dissolved in 0.5 ml of the same solvent and applied to the cartridge. The cartridge was washed with the above solvent mix to elute non-bound protein which was collected in a 5 ml fraction. The sulphur-rich albumins were eluted with 40% solvent A–60% solvent B. Both fractions were recovered by lyophilization and 3.5 mg of sulphur-rich albumin fraction was recovered from a single run of the Sep-pak column. SFA 7 and SFA 8 were separated by chromatography as described above.

Electrophoresis. Polyacrylamide gel electrophoresis (PAGE) was carried out in 1.5 mm slabs [7.5% (w/v) polyacrylamide] at pH 4.3 [30] and pH 8.8 [31]. SDS-PAGE was carried out in 1.5 mm slabs [12% (w/v) polyacrylamide] as described in ref. [32] or in the Tricine system [10% (w/v) polyacrylamide] as described in ref. [33]. The gels were stained with Coomassie Brilliant Blue G in 7.5% (v/v) HOAc/50% (v/v) MeOH soln and destained in 7.5% (v/v) HOAc/10% (v/v) MeOH. The M_rs of the albumins were estimated by SDS-PAGE using Pharmacia LMW and PMW standards. The apparent M_rs of the PMW standards were those reported in ref. [33].

Amino acid analysis. Amino acid compositions of the proteins were determined as described previously [29].

Amino-terminal sequence analysis. The amino-terminal sequences of SFA 7 and SFA 8 were determined in an Applied Biosystems sequencer using 1 nmol protein and the phenylthiohydantoin (PTH) derivatives were identified by HPLC.

Spectral analysis. Spectra of the albumins were determined with a Hewlett Packard 1040A photodiode-array detector controlled by an 85B microcomputer during analytical RP-HPLC of the albumins on a Vydac 218TP54 column using a Perkin Elmer series 4 solvent delivery system equipped with the photodiode-array detector. The ratio of the A at 275 to 290 nm indicated the presence of Tyr and/or Trp residues in the proteins.

Acknowledgements—We thank Mr N. Bartone for amino acid analyses and Mr P. Strike for automated amino-terminal sequence analyses.

REFERENCES

1. Higgins, T. J. V. (1984) *Ann. Rev. Plant Physiol.* 35, 191.
2. Derbyshire, E., Wright, D. J. and Boulter, D. (1976) *Phytochemistry* 15, 3.
3. Youle, R. J. and Huang, A. H. C. (1981) *Am. J. Botany* 68, 44.
4. Spies, J. R. (1974) *J. Agric. Food Chem.* 22, 30.
5. Youle, R. J. and Huang, A. H. C. (1979) *J. Agric. Food Chem.* 27, 500.
6. Li, S. S.-L., Lin, T. T.-S. and Forde, M. D. (1977) *Biochim. Biophys. Acta* 492, 364.
7. Lonnerdal, B. and Janson, J. C. (1972) *Biochim. Biophys. Acta* 278, 175.
8. Menendez-Arias, L., Monsalve, R. I., Gavilanes, J. G. and Rodriguez, R. (1987) *Int. J. Biochem.* 19, 899.
9. Sun, S. S. M., Leung, F. W. and Tomic, J. C. (1987) *J. Agric. Food Chem.* 35, 232.
10. Sharief, F. and Li, S. S.-L. (1982) *J. Biol. Chem.* 257, 14 753.
11. Ericson, M. L., Rodin, J., Lenman, M., Glimelius, K., Josefsson, L.-G. and Rask, L. (1986) *J. Biol. Chem.* 261, 14 576.
12. Menendez-Arias, L., Moneo, I., Dominguez, J. and Rodriguez, R. (1988) *Eur. J. Biochem.* 177, 159.
13. Ampe, C., Van Damme, J., de Castro, L. A. B., Sampaio, M. J. A. M., Van Montagu, M. and Vandekerckhove, J. (1986) *Eur. J. Biochem.* 159, 597.
14. Kries, M., Forde, B. G., Rahman, S., Miflin, B. J. and Shewry, P. R. (1985) *J. Mol. Biol.* 183, 499.
15. Joubert, F. J. (1955) *Biochim. Biophys. Acta* 16, 520.
16. Sabir, M. A., Sosulski, F. W. and Mackenzie, S. L. (1973) *J. Agric. Food Chem.* 21, 988.
17. Schwenke, K. D., Schultz, K. J., Linow, K. J., Uhlig, J. and Franzke, Cl. (1974) *Nahrung* 18, 709.
18. Rahma, E. H. and Narasinga Rao, M. S. (1981) *J. Agric. Food Chem.* 29, 518.
19. Prakash, V. and Narasinga Rao, M. S. (1986) *CRC Crit. Rev. Biochem.* 20, 264.
20. Schwenke, K. D. and Raab, B. (1973) *Die Nahrung* 17, 373.
21. Schwenke, K. D., Raab, B., Uhlig, J., Tkocz, H., Behlke, J., Boeltger, M. and Freimuth, U. (1973) *Die Nahrung* 17, 791.
22. Lilley, G. G., Caldwell, J. B., Higgins, T. J., Spencer, D. and Kortt, A. A. (1989) in *Proceedings of the World Congress on Vegetable Protein Utilization in Human Foods and Animal Feedstuffs* (Applewhite, T. H., ed.), pp. 497–502. American Oil Chemists' Society, Champaign, Illinois.
23. Phillips, R. L. and McClure, B. A. (1985) *Cereal Chem.* 62, 213.

24. Marks, M. D., Lindell, J. S. and Larkins, B. A. (1985) *J. Biol. Chem.* 260, 16451.
25. Kirihara, J. A., Petri, J. B. and Messing, J. (1988) *Gene* 71, 359.
26. Higgins, T. J. V., O'Brien, P. A., Spencer, D., Schroeder, H. E., Dove, H. and Freer, M. (1989) in *The Biology of Wool and Hair*, (Rogers, G. E., Reis, P. J., Ward, K. A. and Marshall, R. C., eds), pp. 441–445. Chapman & Hall, London.
27. Reis, P. J. and Schinckel, O. G. (1963) *Aust. J. Biol. Sci.* 16, 218.
28. Spencer, D., Higgins, T. J. V., Freer, M., Dove, H. and Coombe, J. B. (1988) *Br. J. Nutr.* 60, 241.
29. Kortt, A. A. and Caldwell, J. B. (1990) *Phytochemistry* 29 (in press).
30. Reisfeld, R. A., Lewis, U. J. and Williams, D. E. (1962) *Nature* 195, 281.
31. Davis, B. J. (1964) *Ann. N.Y. Acad. Sci.* 122, 404.
32. Laemmli, U. K. (1970) *Nature* 227, 680.
33. Schagger, H. and von Jagow, G. (1987) *Anal. Biochem.* 166, 368.